

Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat

Borrill, Philippa; Harrington, Sophie A; Uauy, Cristobal

DOI:

[10.1111/tpj.14150](https://doi.org/10.1111/tpj.14150)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Borrill, P, Harrington, SA & Uauy, C 2018, 'Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat', *The Plant journal : for cell and molecular biology*. <https://doi.org/10.1111/tpj.14150>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Checked for eligibility 17/12/2018

"This is the peer reviewed version of the following article: Borrill et al., Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. The Plant Journal, , which has been published in final form at <https://doi.org/10.1111/tpj.14150>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions."

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

DR PHILIPPA BORRILL (Orcid ID : 0000-0002-7623-8256)

PROFESSOR CRISTOBAL UAUY (Orcid ID : 0000-0002-9814-1770)

Article type : SI Genome to Phenome

Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat

Philippa Borrill¹*, Sophie A. Harrington²*, Cristobal Uauy²

¹ School of Biosciences, The University of Birmingham, Birmingham, B15 2TT, United Kingdom

² John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

* PB and SAH should be considered joint first authors.

Corresponding author: Cristobal Uauy (cristobal.uauy@jic.ac.uk)

ORCID:

Philippa Borrill: <http://orcid.org/0000-0002-7623-8256>

Sophie A. Harrington: <http://orcid.org/0000-0003-0754-0678>

Cristobal Uauy: <http://orcid.org/0000-0002-9814-1770>

Running title: Trait discovery in polyploid wheat

Keywords: Genetics, phenotyping, crop improvement, *Triticum aestivum*, natural variation, homoeolog, polyploidy, genetic diversity, gene validation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.14150

This article is protected by copyright. All rights reserved.

Summary

Improving traits in wheat has historically been challenging due to its large and polyploid genome, limited genetic diversity, and in-field phenotyping constraints. However, within recent years many of these barriers have been lowered. The availability of a chromosome-level assembly of the wheat genome now facilitates a step-change in wheat genetics and provides a common platform for resources including variation data, gene expression data, and genetic markers. The development of sequenced mutant populations and gene-editing techniques now enable the rapid assessment of gene function in wheat directly. The ability to alter gene function in a targeted manner will unmask the effects of homoeolog redundancy and allow the hidden potential of this polyploid genome to be discovered. New techniques to identify and exploit the genetic diversity within wheat wild relatives now enable wheat breeders to take advantage of these additional sources of variation to address challenges facing food production. Finally, advances in phenomics have unlocked rapid screening of populations for many traits of interest both in greenhouses and in the field. Looking forwards, integrating diverse data types, including genomic, epigenetic and phenomics data will take advantage of big data approaches including machine learning to understand trait biology in wheat in unprecedented detail.

Introduction

Crop production must increase to meet the food, feed, and fuel demands of a global population estimated to exceed nine billion by 2050 (UN, 2015). Currently, one in nine people live under food insecurity (FAO, 2015). With limited opportunity to expand agriculture on existing land, increasing yields could significantly reduce the number of people at risk of hunger (Rosegrant et al., 2013). Despite the need for a 50% increase in crop production by 2050 (Tilman et al., 2011), our current rates of yield increase are insufficient to reach this goal (Ray et al., 2013). It is therefore critical and urgent that we identify approaches to increase crop productivity, for example through genetically improving cultivars (see Box 1 for glossary of terminology) and improving agronomic practises (Hawkesford et al., 2013, Spiertz, 2012).

Here we focus on polyploid wheat, which accounts for more than 20% of the protein and caloric intake of humans. Compared to other major (diploid) cereals, e.g. rice and maize, wheat has lagged in the development of genomic resources and the genetic understanding for many major productivity traits (e.g. yield, abiotic stress tolerance). This lag was principally due to its polyploid genome which makes genetic analyses in wheat more cumbersome than in diploids (Bevan et al., 2017) and is often a barrier for researchers to understand the biology directly underlying traits in wheat (Box 1). The fact that many productivity traits are inherited in a quantitative manner and require field-based phenotyping also compounds the issues. However, over the past few years, many of the entry barriers to wheat research have been dramatically lowered thanks to multiple genomics developments and the availability of open access resources (Uauy, 2017). For many purposes, wheat can now be treated (almost) like a model crop species for trait discovery.

The objective of this review is to discuss how these latest developments can be applied to understand the biology of agronomic traits in polyploid wheat. We describe the new resources and technological developments in the field and discuss how they can be integrated to improve gene discovery and validation in wheat. We argue that an improved mechanistic understanding of agronomically relevant traits is required for the most efficient deployment of induced and natural variation into the field. This is especially relevant in the polyploid context, where genetic redundancy acts to conceal phenotypic variation. We hope that this review will encourage a wide range of scientists to undertake research for trait discovery in wheat and to further understand biological mechanisms in this fascinating polyploid species. This is an urgent task for achieving global food and nutritional security given the importance of wheat in human diets.

Loss of diversity during domestication and breeding

On average, each person consumes over 40 wheat plants each day (Supporting Data S1). The majority of this consumption (95%) is in the form of bread wheat (*Triticum aestivum*; AABBDD genomes) which is a hexaploid species and most commonly used for bread-making and biscuit flours. A smaller proportion (5%) is accounted for by durum wheat (*T. durum*;

AABB genomes) which is a tetraploid species consumed as bulgur and also used to make pasta and couscous. Although distinct species, both share wild tetraploid emmer wheat (*T. dicoccoides*, AABB genome) as their common ancestor (Figure 1). While bread wheat arose by the hybridization of domesticated emmer wheat with diploid goat grass (*Ae. tauschii*, DD genome), pasta wheat arose directly from emmer wheat following two domestication events (Dubcovsky and Dvorak, 2007).

This evolutionary history, in a polyploid context, has multiple consequences which are important to understand trait biology in wheat (Figure 1). Both modern-day pasta and bread wheat have reduced genetic variation (less than 30%) compared to their wild relatives. This is to an extent due to the domestication bottleneck and selection pressure during breeding (Gaut et al., 2018). Even so, this genetic diversity is lower than that found in equivalent diploid cereals, such as maize and millet, which retain over 60% of the genetic diversity from wild relatives (Wright et al., 2005). The loss of diversity in wheat is mostly due to the reproductive isolation between closely related species of different ploidy levels, although this loss is not equivalent for all genomes. For instance, the A and B genomes of modern bread wheat retain roughly 30% of the genetic diversity present in the wild emmer progenitor (Haudry et al., 2007), whereas the D genome captures less than 10% of the genetic diversity found in goat grass (Halloran et al., 2008). This is because inter-breeding can occur between wild tetraploid emmer (AABB) and hexaploid wheat, whereas diploid goat grass (DD) and species of higher ploidy levels (such as tetraploid and hexaploid wheat) are reproductively isolated (Dvorak et al., 2006, Sutherland and Galloway, 2016)(Figure1). This reduced diversity in modern wheat argues strongly for the use of induced and natural variation to provide novel genetic variation which has been hitherto unexplored in modern breeding programmes (Winfield et al., 2017).

Homeologs and buffering

The A, B, and D genomes have a set of complementary genes (referred to as homoeologs, Box 1) which share between 95-98% sequence identity across coding regions (IWGSC et al., 2018). As a result, many wheat genes are present as two or three functional homoeologous copies (Borrill et al., 2015). This affects the phenotypic consequences of variation at a single

locus, since this can be masked by redundant copies on the other homoeologous genomes. This is especially relevant for quantitative traits where the genetic complexity of the phenotype is affected not only by the multiple genes that control the trait, but also by the multiple homoeologs of each of those genes. This two-fold complexity makes the study of quantitative traits particularly challenging in polyploid species. The main effects of polyploidy on phenotypes are described below and represented in Figure 2:

a) Dominance: Mutations in a single homoeolog lead to a gain-of-function allele which effectively dominates over the other homoeologous copies. Examples include genes with natural variation in non-coding regions such as vernalisation (*VRN1*, Yan et al. (2003)) and photoperiod response (*PPD*, Wilhelm et al. (2009)).

b) Dosage effect: Changes in one homoeolog lead to subtle phenotypic effects, which are amplified by combining variation in additional gene copies. This dosage effect can be (i) additive where mutations in increasing numbers of copies lead to proportionally more extreme phenotypes or (ii) non-additive whereby single or double mutants have subtle phenotypes which become more extreme when combined as triple mutants. Additive examples include genes controlling grain protein content (*GPC*, Avni et al. (2014)), grain size (*GW2*, Wang et al. (2018b)) and red pericarp colour (*R*, Himi et al. (2011)), whereas non-additive genes include *Waxy* controlling amylopectin content (Kim et al. (2003)).

c) Functional redundancy: Mutations in one or two homoeologs of the gene lead to phenotypes which are indistinguishable from wild type plants with three functional versions. Only when all three copies are manipulated simultaneously is there a significant difference in phenotype. Examples include genes controlling important quality traits such as amylose content (*SBE-II*, Slade et al. (2012)), and recessive disease resistance (*MLO*, Acevedo-Garcia et al. (2016)).

During domestication and breeding, humans have selected most frequently for dominant gain-of-function mutations given their strong, easily identifiable phenotypic effects. These include genes with major effects on phenology and adaptation traits as described above, as well as most disease resistance genes (Uauy et al., 2017). In contrast, the effect of recessive

loss-of-function mutations in a single homoeolog are frequently masked by functional redundancy and dosage effects (Figure 2). Hence, phenotypic variation conferred by loss-of-function mutations in a single homoeolog are almost indistinguishable from wild type plants. This makes their identification difficult and consequently these loci are seldomly detected. In some cases, the variation caused by the loss of a single homoeolog is sufficient to allow their detection as quantitative trait loci (QTL), although the molecular cloning of genes underlying this subtle variation in polyploid wheat has been extremely difficult to date (e.g. *GPC*). This also explains why many of the examples outlined above (*SBE-II*, *MLO*, *Waxy*, *GW2*, etc) were not identified in wheat through phenotypic variation, but rather were characterised in wheat as candidate genes based on knowledge from diploid model species. This leads us to conclude that the wheat genome holds a huge potential of functional variation that has remained hidden and under-exploited to date. We argue that gene discovery in wheat is especially relevant to fully exploit this hidden variation as it enables the combination of allelic variation across homoeologs which in many cases is required to expand the phenotypic spectrum. Identification of causal loci also allows a deeper understanding of allelic variation enabling its fine-tuning to modulate the plant's phenotype in different environments (reviewed in Borrill et al. (2015)).

Laying the foundations for understanding genetic diversity

Genome Assemblies

Attempts to sequence the wheat genome have been hindered by its large size (16 Gb), polyploid nature, and large proportion of repetitive sequences (>85%). Despite this, since 2014 several genome assemblies have been produced for the reference landrace Chinese Spring including the International Wheat Genome Sequencing Consortium (IWGSC) Chinese Spring Survey Sequence (IWGSC, 2014) and the TGACv1 assembly (Clavijo et al., 2017). These two assemblies had good coverage of genic regions but suffered from fragmentation principally in repetitive regions. More recently, the use of long read technology enabled the production of a more contiguous sequence, the Triticum 3.0 reference (Zimin et al., 2017) which assembled in total 15.3 Gb. However, this genome was not annotated with gene models which restricts its use for gene function studies.

Recently, the IWGSC RefSeqv1.0 genome was released (IWGSC et al., 2018), providing an annotated chromosome-level assembly for hexaploid wheat. RefSeqv1.0 has a super-scaffold N50 of 22.8 Mb and a total assembly size of 14.5 Gb, representing 94 % of the whole genome. In total 107,891 high-confidence gene models were annotated using Illumina and PacBio reads, providing a stable nomenclature from which to build upon. The availability of these gene annotations and the high sequence contiguity have encouraged the uptake of the RefSeq1.0 genome as a community standard, now available through *EnsemblPlants* (Bolser et al., 2017). Multiple datasets have now been incorporated into this common reference such as the physical position of over 4.7 M genetic markers, over 10 M chemical induced TILLING mutants (Krasileva et al., 2017), and over 850 RNA-Seq datasets (Ramírez-González et al., 2018).

The availability of chromosome-level assemblies is facilitating a step-change in wheat genetics, enabling the rapid correlation of genetic and physical maps. Whereas historically gene cloning in wheat has relied almost entirely upon genetic mapping of traits, these new alignments will now accelerate trait discovery by placing genetic markers in a physical context and allowing rapid identification of candidate genes within mapped regions. A limitation, however, is that in some cases the population used for genetic mapping may have a different gene content than the reference sequence cultivar, which could lead to candidate genes being missed. This drawback is being addressed through pangenome efforts.

Pangenomes

With the completion of a high-quality reference sequence for the wheat landrace Chinese Spring, projects are now focusing on sequencing different wheat cultivars. This is not only important to understand variation within breeding pools and diversity collections, but also to identify sequences and gene content which may be different between the single reference accession and other cultivars used in mapping populations.

Initial efforts focused on the diploid and tetraploid progenitors of wheat: high-quality assemblies for the known A and D genome diploid progenitor species, *T. urartu* and *Ae. tauschii*, respectively, alongside the AABB tetraploid progenitor *T. dicoccoides*, were released in the past 24 months (Ling et al., 2018, Zhao et al., 2017, Avni et al., 2017). A high-

quality reference genome for tetraploid durum wheat, based on the cultivar Svevo is also available (<https://www.interomics.eu/durum-wheat-genome>). Concurrently, the 10+ Wheat Genomes project was launched with the aim to provide high quality assemblies for cultivars valuable for both research and breeding (www.10wheatgenomes.com). The project aims to sequence 18 different bread wheat cultivars from around the world, some of which are already accessible via online BLAST servers (<https://wheatis.tgac.ac.uk/grassroots-portal/blast> and http://webblast.ipk-gatersleben.de/wheat_ten_genomes/). The availability of highly contiguous reference genomes for multiple cultivars of relevance in elite wheat breeding programmes will be particularly important to identify structural variants and copy number variation. This type of variation is known to affect traits of agronomic relevance (Gaut et al., 2018), yet is difficult to detect with traditional genotyping platforms.

Genotyping

Alongside the advances in genomic resources, there is a need to rapidly and accurately assess the variation in specific germplasm. The use of “SNP-CHIPS”, arrays of thousands of cultivar SNPs (single nucleotide polymorphisms), has rapidly increased the amount and accessibility of genotype data for specific cultivars of interest (reviewed in (Uauy, 2017)). As sequencing methods and imputation methods improve (Wang et al., 2018a), low coverage (skim) sequencing will soon be a viable option in polyploid wheat (Golicz et al., 2015).

Alongside the large-scale SNP datasets available at CerealsDB (Wilkinson et al., 2016) and *EnsemblPlants* (Bolser et al., 2017), advances in PCR genotyping markers have facilitated moves into marker-assisted selection in breeding programs. In particular, the use of the KASP (Kompetitive allele specific PCR) technology has provided a sensitive method for bespoke markers within a polyploid context (Ramirez-Gonzalez et al., 2015). Pre-designed KASP markers are available from CerealsDB and MASWheat for agronomically-relevant SNPs such as *GW2*, a grain size gene which is being used in breeding programs (Simmonds et al., 2016). These genotyping methods have demonstrated that the polyploid nature of wheat need not be a barrier to obtaining high-quality, high-throughput genetic variation information.

Use of natural variation for trait discovery

The evolutionary history of polyploid wheat offers several potential sources of genetic variation which can be accessed with different degrees of ease. As discussed above, modern polyploid wheat cultivars have lower genetic variation incorporated from their wild relatives than other major cereal crops such as maize and millet. This lack of diversity is exacerbated for the D genome of bread wheat as only a few *Ae. tauschii* accessions (D-genome progenitor) are thought to have been introduced into hexaploid wheat during the natural hybridization process. Thus, efforts have focused on reintroducing variation from *Ae. tauschii* through the re-synthesis of bread wheat. This is achieved by crossing a tetraploid wheat with a diploid *Ae. tauschii* accession, giving rise to a synthetic hexaploid wheat (SHW) (Dreisigacker et al., 2008). These SHW have been shown to be valuable sources of resistance to abiotic and biotic stress (e.g. Cossani and Reynolds (2015)). Importantly the introduced D-genome from the SHW can freely recombine with modern hexaploid wheat and thus traits can be easily transferred into modern cultivars.

Similar efforts have been made to introduce variation from both evolutionarily near and distant wild wheat relatives (King et al., 2016), although these strategies require specialised genetic stocks to allow the transfer of wild wheat chromosome segments. These segments do not freely recombine, therefore, in addition to the beneficial gene, large DNA fragments from the wild relative are introduced often with negative consequences (linkage drag). The use of *ZIP4* mutants can be used to circumvent some of these limitations (Rey et al., 2017). In an era of gene-editing and in which most wild wheat relatives are sequenced (Ling et al., 2018, Zhao et al., 2017, Avni et al., 2017), it may be more efficient to identify genetic variation of interest for a given trait through studies directly in the wild relatives (Golan et al., 2015). This genetic variation can then be reproduced or inserted into polyploid wheat via gene-editing, circumventing the technically difficult introgression process and potential linkage drag. Initial efforts in this direction are already proving useful. Arora and colleagues phenotyped 174 *Ae. tauschii* accessions with several isolates of the stem rust pathogen and sequenced the disease resistance genes from these accessions (Arora et al., 2018). Performing a k-mer based association approach called AgRenSeq, they cloned four dominant stem rust resistance genes within months. This knowledge can now be used to

introduce the beneficial alleles via transformation, mine for these alleles within modern hexaploid cultivars, or to modify modern sequences via gene-editing.

An additional important source of accessible genetic variation is found in landraces, locally-adapted wheat lines which were grown as farmer-saved seed and that have not been modified through modern breeding techniques. These landraces constitute reservoirs of genetic diversity that can be directly cross-bred into modern cultivars. There are thousands of these landrace accessions available, however many of them are heterogeneous, often duplicated seed stocks (Singh et al., 2018), which remain confined to germplasm repositories. Of special note is the Watkins landrace collection (Wingen et al., 2014) encompassing over 1,000 landraces collected in the 1920's and 1930's from a wide geographic distribution. This collection has been purified by single-seed descent from which many genomic and genetic resources have been developed. Genotyping of this collection revealed extensive novel genetic (Winfield et al., 2017) and epigenetic (Gardiner et al., 2018b) diversity which was absent from modern cultivars. A core set of 107 accessions were used to generate nested association mapping populations (Wingen et al., 2017), all of which were genotyped, have genetic maps available, and are free to access (<http://wisplandracespillar.jic.ac.uk/>).

Modern cultivars also provide useful genetic variation for trait discovery despite this variation being overall reduced with respect to landraces and wild relatives (Figure 1). Much of this variation is studied through genetic populations composed of cultivars which differ for the trait(s) of interest. These include bi-parental populations between two cultivars (e. g. Saintenac et al. (2018)), multi-parent advanced generation inter-cross (MAGIC) populations composed of between 4 to 16 cultivars (Huang et al., 2012, Mackay et al., 2014, Dixon et al., 2018, Milner et al., 2015), nested association mapping panels of multiple cultivars to a common parent (Jordan et al., 2018), and association panels of 100 or more cultivars (Liu et al., 2017, Sukumaran et al., 2015). All of these types of populations are available in wheat and their relative merits are discussed elsewhere (Huang and Han, 2014). Importantly, the coupling of these populations to the new genomic and validation tools (discussed below) have provided an integrated workflow for gene discovery in wheat, exemplified by the Triticeae Toolbox from the TCAP project (Blake et al., 2016).

The availability of phenotypic data for these vast collections of genotyped landraces, cultivars and wild relatives opens the possibility to assign a breeding value to each polymorphism. This approach enables the implementation of statistical models to define the best combinations of SNPs to maximise the breeding value of progeny from a genetic cross. This genomic selection approach, first implemented in cattle, is used in many crops including wheat (Bassi et al., 2016). This approach, however, does not aim to perform trait discovery as it is agnostic of the functional effect of the polymorphism being selected. While effective for highly heritable traits and major effect loci, many of which are already known, this approach is best suited for standing variation within modern cultivars and makes it difficult to incorporate diversity from non-adapted cultivars or accessions.

Moving from SNPs to haplotypes in breeding

Breeders are constantly shuffling SNPs by making specific crosses and exploiting the subsequent recombination that occurs during meiosis. These SNPs, however, are not inherited independently, but rather as haplotypes; blocks of polymorphisms with a defined order and which in many cases extend across several genes. As genome assemblies have improved in wheat, we are now able to anchor the millions of SNPs previously identified onto the full reference sequence. With this data, it becomes clear that, in the absence of the causal polymorphism *per se*, it is important to recognise the haplotype as the unit of genetic information that underlies traits, rather than a closely associated bi-allelic SNP. The case study of the wheat pre-harvest sprouting resistance gene *TaMKK3* provides a telling example to this effect (Barrero et al., 2015, Torada et al., 2016, Shorinola et al., 2017). Thus, when making crosses, breeders are shuffling haplotypes to generate new combinations of these blocks of DNA.

The identification of haplotypes allows us to investigate and improve traits both through a retrospective and a prospective approach (Bevan et al., 2017). The retrospective approach takes advantage of the rich history of wheat breeding which has very well-defined pedigrees for most cultivars, many of which can be traced back to landraces. Given that recombination is not evenly distributed across the length of the chromosomes (Akhunov et al., 2003), many haplotypes in elite cultivars today will have remained intact as a unit during this breeding

process (albeit shuffled and assembled in different combinations with other haplotypes). This allows us to retrospectively evaluate which haplotypes have been selected and enriched for in certain environments and which haplotypes have been lost, potentially due to deleterious effects. Thus, we can benefit from years of available phenotypic data and selection to assess haplotype x environment interactions. This retrospective approach has highlighted, for example, how strong selection for flowering time has eliminated haplotypes for increased root biomass, a desirable trait (Voss-Fels et al., 2016). We can now use this knowledge to break these negative correlations by purposely generating haplotypes that combine both beneficial traits. This will be facilitated by ongoing studies into the genetic basis of recombination in wheat which could provide the knowledge to manipulate recombination rate distribution across chromosomes (Jordan et al., 2018). With clearly defined haplotypes, we can also leverage the multiple layers of genomic information (e.g. networks, RNA-Seq; Figure 3) to help identify the causal variant underlying the trait of interest.

By understanding the haplotypes available within elite cultivars, we can also use a prospective approach to mine landrace and wild relative collections for novel haplotypes. This approach could entail mining for haplotypes which break negative correlations across specific low-recombinogenic regions (Winfield et al., 2015), or alternatively mining for novel haplotypes across QTL loci associated with a known agronomic trait (Muterko et al., 2015). This will become more commonplace as large diversity collections are genotyped at high-density and haplotype imputation becomes more robust (Jordan et al., 2015). In this manner, rather than phenotyping a large diversity collection for many traits, the phenotypic value of different haplotype blocks across a specific interval could be examined in further detail. The Watkins nested association mapping populations offer a valuable resource to this effect as the landrace haplotypes can be directly evaluated within available mapping populations (Wingen et al., 2017). As breeding programs increasingly look to combine genotypic and phenotypic data, the availability of haplotype data for wild relatives, landrace collections, and elite wheat cultivars with their associated phenotypic value, has the potential to shape the way in which breeding programs are structured.

New strategies accelerate gene discovery in wheat

With the genomic resources developed recently for wheat, many new techniques for gene discovery have been developed. One area has focused on rapid cloning of disease resistance (*R*-) genes by using targeted genome sequencing approaches (Keller et al., 2016, Periyannan, 2017). MutRenSeq takes advantage of significantly improved *R*-gene annotations to use an *R*-gene enriched capture method to sequence the majority of *R*-genes in a given plant (Jupe et al., 2013, Steuernagel et al., 2016)(Figure 3). Combining *R*-gene enrichment sequencing with screening of mutagenized populations for dominant disease resistance phenotypes has allowed easier and more rapid cloning of *R*-genes than previously possible via traditional genetic mapping (Marchal et al., 2018). As described above, a variation on the technique, AgRenSeq, uses the enrichment strategy to identify variation in the *R*-genes of wild relatives (Arora et al., 2018).

While the whole genome sequence is now available, it remains cost and time-prohibitive in most cases to fully re-sequence lines for gene discovery. Instead, integrating known genetic mapping with new genomic approaches can accelerate trait discovery. Exome-capture sequencing has been used to identify variation in the coding region of the wheat genome (Saintenac et al., 2011) and can be used to identify candidate genes, often in combination with bulk-segregant analysis (Gardiner et al., 2016, Mo et al., 2018). SNPs identified from the exome capture can be used to both develop new markers for genetic mapping and to identify mutations in putative candidate genes. However, this technique is limited by the quality of the capture probes used. Many current capture arrays were designed based on the gene models from the 2014 genome, and as such are missing or mis-representing various genes (i.e. Brinton et al. (2018)). Efforts are underway to develop new capture arrays based on the improved annotations (Gardiner et al., 2018a).

Alternative techniques aim to reduce the size of the genome by flow-sorting and sequencing single chromosomes (Giorgi et al., 2013). One approach, MutChromSeq (Sánchez-Martín et al., 2016), requires a mutation-induced phenotype which can be readily-screened and has been mapped to a specific chromosome. The target chromosomes of independent mutants are then flow-sorted and sequenced before being compared to identify candidate genes. A second approach, TACCA (Thind et al., 2017), relies on knowledge of a specific target

interval (based on flanking markers) identified by traditional genetic mapping, and as such is amenable to both qualitative and quantitative traits (Box 1). The aim of this method is to build, in one step, long range contiguous assemblies that completely span the target interval, thereby eliminating the need for BAC clones. Flow-sorted chromosomes from the cultivar of interest are sequenced and assembled using the Chicago method. If the resulting long-range assemblies are contiguous enough to contain the flanking genetic markers, then identification of the causal variation can be rapid. These approaches allow for the un-biased discovery of variation beyond the genic variation identified with exome capture and also aids the identification of causal variants in cultivars that differ from the established reference cultivars. Looking forward, CRISPR-Cas9 may be used to cut at flanking marker sites to generate a relatively short piece of genomic DNA which can be separated from other genomic DNA by size selection before sequencing. This approach which has been carried out using human DNA (Nachmanson et al., 2018), albeit with relatively closely spaced cut sites, would avoid the requirement for chromosome flow-sorting which is technically challenging and is only routinely carried out in a few laboratories worldwide.

Hypothesis generation in wheat

Orthology

One of the more tangible results of the improved genome assemblies and gene annotations is the increased ease with which knowledge from other species can be transferred into wheat (Adamski et al., 2018). High quality gene annotations based on the new assemblies are now included in platforms such as *EnsemblPlants* which identify putative orthologous genes in other species. This knowledge base facilitates the application of reverse genetic approaches to wheat using orthologs of genes identified in other species. Once identified, these genes can be tested for function in wheat using sequenced mutant populations (Krasileva et al., 2017), as will be discussed in more detail below. An important caveat to any reverse genetics approach is that known function cannot be presumed to be conserved between species, particularly when moving from a dicot species such as *Arabidopsis* into wheat. Comparative studies on the flowering time pathways between *Arabidopsis* and cereals show that while some core pathways have been conserved across species, there is

lineage-specific divergence (Higgins et al., 2010). However, the candidate gene strategy has been successful in identifying putative genes in wheat, particularly from rice orthologs, which have gone on to be characterised and confirmed, such as *GW2* (Simmonds et al., 2016, Zhang et al., 2018, Wang et al., 2018b) and other regulators of grain size (Li and Yang, 2017).

Developments of resources in other Triticeae species, especially barley, may also enable the rapid transfer of knowledge amongst this phylogenetic clade. This approach has been successful for genes relating to e.g. flowering time (Beales et al., 2007, Turner et al., 2005), cuticular wax biosynthesis (Hen-Avivi et al., 2016, Schneider et al., 2016) and pre-harvest sprouting (Nakamura et al., 2016, Torada et al., 2016). Likewise, genes identified in tetraploid wheat can be rapidly evaluated in hexaploid wheat as most genes identified in emmer and durum wheat perform equivalent roles in hexaploid wheat (Uauy et al., 2005, Uauy et al., 2006).

Expression

The decreasing cost of RNA-Seq and improved wheat genome sequences has enabled a rapid increase in large-scale gene expression analyses in wheat (Figure 3). Software, such as the pseudoaligner kallisto (Bray et al., 2016) can readily distinguish between sequences from highly similar homoeologs (Borrill et al., 2016), allowing homoeolog-specific gene expression analyses. These analyses have been used to understand the pathways regulating varied developmental and stress-related traits (Kugler et al., 2013, Brinton et al., 2018), to understand the roles of individual genes (Pearce et al., 2014, Pearce et al., 2016), and to compare response between stress treatments (Zhang et al., 2014, Liu et al., 2015). These examples highlight the many ways in which RNA-Seq datasets can be used to shed light on the regulatory pathways involved in different traits of interest and to identify candidate genes for hypothesis generation.

To maximise the utility of this data many of these studies have been incorporated into gene expression atlases. The expVIP gene expression atlas contains over 900 RNA-Seq samples from a range of diverse tissues, developmental stages, stresses (abiotic and disease), and cultivars (Borrill et al., 2016, Ramírez-González et al., 2018). Metadata categorising the samples allows for comparisons between different studies and enables sorting and filtering

of relevant samples on the web interface (www.wheat-expression.com). A highly detailed developmental timecourse of the spring wheat cultivar Azhurnaya (Ramírez-González et al., 2018) is available both through expVIP and a dedicated eFP browser (http://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi) (Winter et al., 2007). This pictographic display facilitates rapid assessment of the expression pattern of a gene across 70 different tissues and developmental stages.

This readily accessible gene expression data enables hypothesis generation based on expression patterns (Figure 4). Comparisons of expression data between orthologous genes can provide insight into shared or divergent functions between species. Expression data can also be a useful source of supporting information to narrow down candidate genes identified through conventional QTL mapping or functional genomic approaches using natural and induced variation.

Networks

Although knowledge about individual gene's expression patterns is very useful, it is also important to consider how genes act together in networks to regulate phenotypes. In model species such as Arabidopsis, network approaches have been successful in identifying candidate genes regulating specific processes (Bassel et al., 2012) and recent comparative work has uncovered conserved gene networks in root tips across monocots and dicots (Maher et al., 2018). Network-enabled approaches are likely to find broad applications in wheat research, now that the necessary genomic resources are available (Figure 3).

Early work by WheatNet (<https://www.inetbio.org/wheatnet>) generated a wheat network using 20 genomics datasets, although only one of these was derived from wheat specific data (Lee et al., 2017). Knetminer (<http://knetminer.rothamsted.ac.uk>) also presents a wheat network, which until recently was based mainly on information derived from orthology (Hassani-Pak et al., 2016). The use of co-expression studies has been relatively limited, although genes and pathways associated with fusarium head blight resistance (Kugler et al., 2013) and spike architecture (Wang et al., 2017) have been uncovered.

Recently, a set of co-expression networks were developed using publicly available and newly generated RNA-Seq samples for specific tissue types (leaf, root, grain and spike) and abiotic or disease stress conditions (Ramírez-González et al., 2018). The development of an

integrated network for all 850 RNA-Seq samples provides an entry point into understanding the general co-expression patterns of genes across a wide range of tissues (IWGSC et al., 2018) and has been incorporated into Knetminer. Expression data has also been leveraged to generate a network which predicts downstream targets of transcription factors (Ramírez-González et al., 2018), also available in Knetminer. A powerful approach in the future will be to layer wheat-specific information on top of information from model species to identify candidate genes and groups of genes regulating specific traits of interest.

An important question to consider when working with a polyploid species is whether the homoeologous copies of genes are functionally redundant. These co-expression networks have started to address this question and show that whilst many homoeologs have the same or similar expression patterns, approximately 30 % of homoeologs have divergent expression patterns from each other (Ramírez-González et al., 2018). This implies that a significant proportion of wheat genes may already be undergoing the first steps towards sub/neo-functionalisation through changes in their patterns of expression. This information can be used to inform strategies to alter gene function for example altering a gene whose homoeolog has a very divergent expression pattern is more likely to cause a phenotypic effect than if the homoeologs had highly similar expression patterns.

Gene validation in polyploid wheat

As the new genome assemblies increase the potential for gene discovery, the focus now turns to ways in which gene function can be characterised (Figure 4). Exome-capture followed by next-generation sequencing has been applied to tetraploid and hexaploid TILLING populations to identify and make available induced variation in gene coding regions (Krasileva et al., 2017). This database allows for “*in silico* TILLING” whereby mutations in any gene of interest can quickly and easily be identified. Importantly, while the SNPs were originally called against the 2014 IWGSC reference genome (available at www.wheat-tilling.com and <https://dubcovskylab.ucdavis.edu/wheat-tilling>), they have recently been remapped to the RefSeqv1.0 gene models and are available on *EnsemblPlants*. The mutagenized populations are particularly useful due to their ease of access, with researchers being able to identify mutations in their gene of interest within minutes.

However, since the sequenced mutations will lead to a truncation (premature termination codon or splice junction variant) in only 60% of genes, the use of missense mutations will also be required at times. As a result, screening more than one mutation per gene might be necessary as predicting the functional effects of missense mutations is not always easy *a priori*. Similarly, use of the TILLING lines typically requires multiple iterations of crossing, first to combine mutations across homoeologs of the gene of interest and secondly to reduce background mutations (Krasileva et al., 2017, Uauy et al., 2017). Typically, for quantitative traits, phenotypes are clearer with successive backcross generations (Simmonds et al., 2016). However, some qualitative traits can be studied directly in the initial mutant population (Mo et al., 2018, Shorinola et al., 2018). A caveat for the use of these populations is that not all genes will be present in the *in silico* database. This is either a result of the genes not being present in the exome-capture array, or because these genes are not present in the cultivars used to create the populations. In these cases, identifying mutations in the gene(s) of interest would require the screening of the population using conventional TILLING approaches or the development of a new bespoke mutant population using the cultivar with the gene(s) of interest.

Alternative methods for gene functional characterisation focus on transgenic techniques. These can be either transient or stable transformation methods and are able to influence gene expression levels beyond what is possible via mutagenesis. Transient methods benefit from the relative ease and speed of transformation compared to stable integration, with several transient gene expression and silencing systems now available in wheat. The use of virus-induced gene silencing (VIGS) in wheat and other cereals has been highly successful in characterising the impact of various genes involved in abiotic and biotic stress response (Cakir et al., 2010, Ramegowda et al., 2014, Lee et al., 2015). These VIGS systems are now well established for use in wheat, and more recently advances in virus-mediated overexpression (VOX) have demonstrated the feasibility of their use in wheat for both small and large proteins (Cheuk and Houde, 2018, Bouton et al., 2018). Beyond viral expression systems, other possibilities for transient expression in wheat exist. Nanomaterials have recently shown potential for the transient delivery of foreign DNA into wheat leaves and protoplasts (Demirer et al., 2018).

Despite their increased cost and lengthier transformation process with respect to transient systems, traditional stable transgenics continue to be used in many contexts to validate and explore the function of genes in wheat (reviewed in Borrill et al. (2015)). A major bottleneck of current wheat transformation systems is that only very few, typically non-elite and non-adapted wheat cultivars (such as Fielder or Bobwhite) are transformable at reasonable frequencies (Harwood, 2012), although this is changing (Richardson et al., 2014, Ishida et al., 2015, Wang et al., 2016). Studies in maize have demonstrated that the concurrent expression of transcription factors *Baby boom* and *Wuschel2* increase both transformation efficiency and the range of cultivars which are transformable (Lowe et al., 2016). Integration of this system into wheat transformation platforms could allow genotype-independent transformation which would accelerate the timeline for integration of transgenic traits into wheat breeding pipelines, while also making it easier to evaluate the effects of the transgene or modification in a locally relevant cultivar.

Genotype-independent transformation is of particular interest when considering applications for gene editing, which are not regulated as genetically modified organisms in North America, Japan and elsewhere, though they are in Europe. Gene editing systems such as TALENs and CRISPR-Cas9 have now been demonstrated to function in wheat including the production of transgene-free edited lines (Wang et al., 2014, Zhang et al., 2016, Rey et al., 2018). Beyond the most common use of CRISPR-Cas9, to induce small deletions that lead to frame-shifts and knock-outs, this may involve more complex editing such as targeted gene insertion or specific base editing (reviewed in Adli (2018)). In theory, the use of guide-mediated gene editing could produce triple homoeolog knock-outs in a single transgenic event, although efficiencies remain low (Zhang et al., 2016). As the use of this system becomes more commonplace in wheat, we anticipate that advances in transformation and editing efficiency will follow, increasing the impact of the system. Crucially, the ability to transform elite wheat cultivars would significantly accelerate the rate at which edited alleles could be deployed in farmers' fields.

Developments in phenotyping

Historically, wheat phenotyping for major agronomic and quality traits (e.g. yield, protein content) has required field trials which are labour intensive and can only be carried out once per year in many parts of the world. To overcome this bottleneck (Furbank and Tester, 2011) one area of recent growth has been the use of high-throughput phenotyping systems. Protocols have been developed for controlled environments and field conditions, enabling a wide range of traits to be studied. Large collections of germplasm in wheat have been amassed (see above) but as yet relatively few collections have been screened using high throughput phenomics, presenting an opportunity for future research.

Many traits, including growth rates, biomass, and abiotic stress responses, have been studied in high throughput phenotyping platforms in glasshouse conditions, particularly in specifically constructed facilities with conveyor belts for frequent imaging of plants.

Different cameras are required for imaging specific traits: growth rate and biomass can be calculated from standard RGB images (Parent et al., 2015) whereas infra-red imaging has been used to study osmotic stress (Sirault et al., 2009). One area of research which has particularly benefited from glasshouse based phenomics is the understanding of root architecture. Roots are a particularly difficult tissue to study because access in the field requires excavation and cleaning of the soil through methods such as “shovelomics” which has been applied in maize (Trachsel et al., 2011) and recently wheat (York et al., 2018).

Recently developed higher-throughput technologies have enabled the phenotyping of root architecture in wheat seedlings using transparent pots (Richard et al., 2015) or through paper-based screening (Atkinson et al., 2015). These technologies have now been applied to screen sequenced mutation populations to identify mutants with altered root architecture (Shorinola et al., 2018). Adult plant roots pose a greater challenge due to their larger size, but developments in root tracking approaches (Mairhofer et al., 2015, Pfeifer et al., 2015) are increasing the rate at which high resolution root systems can be imaged through magnetic resonance imaging and X-ray microcomputer tomography.

Controlled environment conditions enable the rapid screening of seedling stage plants, however the strength of the correlation between seedling physiological traits and those of plants growing in the field may be poor for some traits. For example, a meta-analysis for

specific leaf area, leaf nitrogen concentration, and yield across a range of crop species only showed a median correlation of 0.26 between the lab and the field (Poorter et al., 2016). However, steps can be taken to narrow this gap, for example by applying abiotic conditions more similar to those found in the field to controlled environment experiments, growing plants in controlled environments in more natural soils, or by planting at densities more comparable to the field (Poorter et al., 2016).

Therefore, in parallel with pot-based approaches in glasshouses, in-field technologies have been developed. These are usually either a fixed platform with a camera which images a small area of selected plots in the field or based on mobile imaging systems such as manually operated platforms or unmanned aerial vehicles. The specifics of these platforms have been reviewed recently (Araus et al. (2018) and references therein) and each offers a unique set of advantages and disadvantages. These platforms all rely on imaging technologies such as RGB and infra-red, mentioned previously, as well as hyperspectral imaging which is particularly popular for in-field imaging. Hyperspectral imaging uses wavelengths across the electromagnetic spectrum to generate indices such as normalized difference vegetative index (NDVI) to measure canopy coverage (Christopher et al., 2016), as well as to calculate models for specific biochemical and physiological traits (Silva-Perez et al., 2018). Hyperspectral imaging can also be used to detect and classify disease and stress symptoms, however collecting appropriate data and developing analysis pipeline can be a challenge within commercial breeding programmes (Lowe et al., 2017). At CIMMYT large scale phenotyping for physiological traits has been incorporated into breeding programmes using both high throughput imaging and conventional phenotyping methods (Reynolds and Langridge, 2016). The phenotypes obtained through these methods have been used to inform physiological breeding, in which lines with complementary physiological phenotypes are crossed. This approach has successfully bred lines with improved phenotypes for a range of traits including yield potential, drought, and heat stress (Reynolds and Langridge, 2016). Subsequent genetic dissection of the physiological traits in the improved lines can reveal the loci underpinning these phenotypes which can then be used in marker assisted selection across the breeding programme. Combining phenomics with genotypic information is also proving useful, for example to improve genomic selection for quantitative traits such as yield (Crain et al., 2018).

Both in-field and glasshouse based phenomics platforms generate huge quantities of image data which requires processing to extract the relevant information. However, in comparison to genomics analyses, pipelines for image processing are far less standardised and represent a major informatics challenge. As these phenomics approaches have become more widely adopted, a number of tools for image processing from small scale glasshouse experiments to large hyperspectral field experiments have been developed (Perez-Sanz et al., 2017). Machine learning are a promising approach to accelerate data processing and improve trait analysis (Singh et al., 2016), although obtaining high quality annotated images will also be essential (Tsaftaris et al., 2016). In the future it is likely that standard procedures for analysis will be developed, analogous to standard genomic pipelines such as GATK for variant calling (McKenna et al., 2010).

Looking forward

Connecting genes to phenotypes in wheat is challenging, but combining the two is urgent and necessary to accelerate breeding efforts to produce enough food for the world's population. Efforts to develop new, adapted wheat cultivars will require improved understanding of the interactions between genotypes, environment, and management practices, in the context of climate change (Challinor et al., 2014). Despite these challenges, there are great opportunities to understand wheat biology in the coming years.

One area of wheat biology which is as yet underexploited is the modulation of individual homoeologs to influence traits. As we increase our understanding of the extent to which homoeologs are functionally redundant, new approaches such as epigenetic modification, altering promoter sequences, or introducing dominant alleles could be exploited to produce a wider range of phenotypes using the genes which wheat already contains. Many of these alterations could build upon novel gene-editing technologies. For example, genome-editing within promoter regions in tomato has expanded the range of fruit sizes, inflorescence branching, and plant architecture (Rodríguez-Leal et al., 2017). Use of this technology in wheat could expand the phenotypic diversity available for key genes. Targeted engineering of individual homoeologs might unlock previously hidden phenotypes, while the ability to

carry out allele replacement will allow the rapid transfer of desirable alleles into elite breeding material.

The impact of chromatin conformation on phenotypic variation is perhaps even less investigated, yet has the potential to provide substantial insights into the connection between genotype and phenotype. In animals and plants, accessible (open) chromatin has been associated with functional non-coding sequences. In maize, for example, <1% of the genome was defined as open chromatin, yet variation within these regions explained 40% of phenotypic variation for many quantitative traits (Rodgers-Melnick et al., 2016), consistent with studies in humans and *Arabidopsis* (Maurano et al., 2012, Sullivan et al., 2014). Techniques such as Hi-C (Belton et al., 2012) and ATAC-seq (Buenrostro et al., 2015) can be applied to the wheat genome to characterise chromatin structure and identify putative regulatory motifs (Figure 3A). As 99% of the wheat genome is non-protein coding, variation in these regions is likely to account for a significant proportion of phenotypic variation. DNA methylation and histone mark variation both within and across cultivars is only beginning to be studied in wheat (Gardiner et al., 2018b, Ramírez-González et al., 2018) and is likely to provide a wealth of information on gene regulation for traits of agronomic importance. Natural epigenetic variation is known to impact on agronomically relevant traits such as fruit ripening and plant height, and was found to underpin artificially selected variation for increased energy use efficiency in canola (reviewed in Gallusci et al., 2017, Springer, 2013). As wheat genetics moves further towards understanding hidden variation derived from polyploidy, it is now also becoming possible to consider untapped variation in the context of chromatin structure and epigenetic modifications.

Advances in phenomics have enabled us to generate large phenotypic datasets, and now opportunities will arise to understand the links between genotype, environment, agronomy, and phenotype at a systematic scale. Machine learning is showing promise in predicting phenotypes in wheat based on genomic data (Grinberg et al., 2018), performing better than classical statistical genetics methods. Likewise crop simulation models based on physiology and phenology of wheat have improved in recent years, yet they currently only incorporate genetic effects relating to relatively basic phenotypic traits (e.g flowering)(Chenu et al., 2017).

The addition of novel data types such as epigenetics and chromatin conformation may contribute to improved wheat phenotypic prediction. Initial studies reveal epigenetic differences amongst landraces (Gardiner et al., 2018b) which are likely to also be observed in elite material. Incorporating diverse data types (genomics, epigenomics, and environmental data) may enable more accurate predictions of phenotypes. The use of computation models will also be useful for the development of networks predicting the phenotypic values of haplotype combinations, epi-alleles, and the outcomes of crosses between cultivars. Current genomic and epigenomic approaches rely on whole tissue samples; moving forward, single-cell approaches (e.g Cao et al. (2017)) will increase resolution in studies of cell-specific process. However, we must be cautious that the ability to simultaneously measure many thousands of genes, or thousands of phenotypes using new technologies may not necessarily improve our ability to understand the underlying biology. Therefore, we believe that combining these new techniques with targeted studies of specific genes and traits promises to provide insights into the core biological processes regulating agronomically relevant traits in wheat.

In summary, the last decade has brought about a revolution in the resources and approaches used for wheat research and breeding. We now have a wide range of resources which provide opportunities to improve our understanding of wheat gene functions, many of which are easily accessible for researchers moving into wheat from other species (Adamski et al., 2018). One of the key challenges moving forwards will be predicting which genes will be most valuable within breeding programmes in diverse and changing environmental conditions.

Acknowledgments

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) through the Designing Future Wheat (BB/P016855/1) and GEN (BB/P013511/1) Institute Strategic Programmes and an Anniversary Future Leaders Fellowship to PB (BB/M014045/1). SAH was supported by the John Innes Foundation. We thank Tobin Florio for graphical design (www.flozbox.com/Science_Illustrated). The authors declare no conflict of interest.

Supporting information

Supporting Data S1: Calculations for daily consumption of wheat

References

- Acevedo-Garcia, J., Spencer, D., Thieron, H., Reinstädler, A., Hammond-Kosack, K., Phillips, A. L. and Panstruga, R. 2016. *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnology Journal*, 15, 367-378.
- Adamski, N. M., Borrill, P., Brinton, J., Harrington, S., Marchal, C. and Uauy, C. 2018. A roadmap for gene functional characterisation in wheat. *PeerJ Preprints*, 6, e26877v1.
- Adli, M. 2018. The CRISPR tool kit for genome editing and beyond. *Nature Communications*, 9, 1911.
- Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L.-L., Echaliier, B., Gill, B. S., Miftahudin, Gustafson, J. P., Lazo, G., Chao, S., Anderson, O. D., Linkiewicz, A. M., Dubcovsky, J., Rota, M. L., Sorrells, M. E., Zhang, D., Nguyen, H. T., Kalavacharla, V., Hossain, K., Kianian, S. F., Peng, J., Lapitan, N. L. V., Gonzalez-Hernandez, J. L., Anderson, J. A., Choi, D.-W., Close, T. J., Dilbirligi, M., Gill, K. S., Walker-Simmons, M. K., Steber, C., McGuire, P. E., Qualset, C. O. and Dvorak, J. 2003. The Organization and Rate of Evolution of Wheat Genomes Are Correlated With Recombination Rates Along Chromosome Arms. *Genome Research*, 13, 753-763.
- Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S. and Cairns, J. E. 2018. Translating High-Throughput Phenotyping into Genetic Gain. *Trends in Plant Science*, 23, 451-466.
- Arora, S., Steuernagel, B., Chandramohan, S., Long, Y., Matny, O., Johnson, R., Enk, J., Periyannan, S., Hatta, M. a. M., Athiyannan, N., Cheema, J., Yu, G., Kangara, N., Ghosh, S., Szabo, L. J., Poland, J., Bariana, H., Jones, J. D. G., Bentley, A. R., Ayliffe, M., Olson, E., Xu, S. S., Steffenson, B. J., Lagudah, E. and Wulff, B. B. H. 2018. Resistance gene discovery and cloning by sequence capture and association genetics. *bioRxiv*, 10.1101/248146.
- Atkinson, J. A., Wingen, L. U., Griffiths, M., Pound, M. P., Gaju, O., Foulkes, M. J., Le Gouis, J., Griffiths, S., Bennett, M. J., King, J. and Wells, D. M. 2015. Phenotyping pipeline reveals major seedling root growth QTL in hexaploid wheat. *Journal of Experimental Botany*, 66, 2283-2292.
- Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., Hale, I., Mascher, M., Spannagl, M., Wiebe, K., Jordan, K. W., Golan, G., Deek, J., Ben-Zvi, B., Ben-Zvi, G., Himmelbach, A., Maclachlan, R. P., Sharpe, A. G., Fritz, A., Ben-David, R., Budak, H., Fahima, T., Korol, A., Faris, J. D., Hernandez, A., Mikel, M. A., Levy, A. A., Steffenson, B., Maccaferri, M., Tuberosa, R., Cattivelli, L., Faccioli, P., Ceriotti, A., Kashkush, K., Pourkheirandish, M., Komatsuda, T., Eilam, T., Sela, H., Sharon, A., Ohad, N., Chamovitz, D. A., Mayer, K. F. X., Stein, N., Ronen, G., Peleg, Z., Pozniak, C. J., Akhunov, E. D. and Distelfeld, A. 2017. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, 357, 93.
- Avni, R., Zhao, R., Pearce, S., Jun, Y., Uauy, C., Tabbita, F., Fahima, T., Slade, A., Dubcovsky, J. and Distelfeld, A. 2014. Functional characterization of *GPC-1* genes in hexaploid wheat. *Planta*, 239, 313-324.
- Barrero, J. M., Cavanagh, C., Verbyla, K. L., Tibbits, J. F. G., Verbyla, A. P., Huang, B. E., Rosewarne, G. M., Stephen, S., Wang, P., Whan, A., Rigault, P., Hayden, M. J. and Gubler, F. 2015. Transcriptomic analysis of wheat near-isogenic lines identifies *PM19-A1* and *A2* as candidates for a major dormancy QTL. *Genome Biology*, 16, 93.
- Bassel, G. W., Gaudinier, A., Brady, S. M., Hennig, L., Rhee, S. Y. and De Smet, I. 2012. Systems Analysis of Plant Functional, Transcriptional, Physical Interaction, and Metabolic Networks. *The Plant Cell*, 24, 3859.

- Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R. and Crossa, J. 2016. Breeding schemes for the implementation of genomic selection in wheat (*Triticum spp.*). *Plant Science*, 242, 23-36.
- Beales, J., Turner, A., Griffiths, S., Snape, J. W. and Laurie, D. A. 2007. A Pseudo-Response Regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 115, 721-733.
- Belton, J.-M., Mccord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y. and Dekker, J. 2012. Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*, 58, 268-276.
- Bevan, M. W., Uauy, C., Wulff, B. B. H., Zhou, J., Krasileva, K. and Clark, M. D. 2017. Genomic innovation for crop improvement. *Nature*, 543, 346.
- Blake, V. C., Birkett, C., Matthews, D. E., Hane, D. L., Bradbury, P. and Jannink, J.-L. 2016. The Triticeae Toolbox: Combining Phenotype and Genotype Data to Advance Small-Grains Breeding. *The Plant Genome*, 9.
- Bolser, D. M., Staines, D. M., Perry, E. and Kersey, P. J. 2017. Ensembl Plants: Integrating Tools for Visualizing, Mining, and Analyzing Plant Genomic Data. In: VAN DIJK, A. D. J. (ed.) *Plant Genomics Databases: Methods and Protocols*. New York, NY: Springer New York.
- Borrill, P., Adamski, N. and Uauy, C. 2015. Genomics as the key to unlocking the polyploid potential of wheat. *New Phytologist*, 208, 1008-1022.
- Borrill, P., Ramirez-Gonzalez, R. and Uauy, C. 2016. expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform. *Plant Physiology*, 170, 2172.
- Bouton, C., King, R. C., Chen, H., Azhakanandam, K., Bieri, S., Hammond-Kosack, K. E. and Kanyuka, K. 2018. Foxtail mosaic virus: A Viral Vector for Protein Expression in Cereals. *Plant Physiology*, 177, 1352.
- Bray, N. L., Pimentel, H., Melsted, P. and Pachter, L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, 34, 525.
- Brinton, J., Simmonds, J. and Uauy, C. 2018. Ubiquitin-related genes are differentially expressed in isogenic lines contrasting for pericarp cell size and grain weight in hexaploid wheat. *BMC Plant Biology*, 18, 22.
- Buenrostro, J. D., Wu, B., Chang, H. Y. and Greenleaf, W. J. 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current Protocols in Molecular Biology*, 109, 21.29.1-21.29.9.
- Cakir, C., Gillespie, M. E. and Scofield, S. R. 2010. Rapid Determination of Gene Function by Virus-induced Gene Silencing in Wheat and Barley. *Crop Science*, 50, S-77-S-84.
- Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C. and Shendure, J. 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, 357, 661.
- Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R. and Chhetri, N. 2014. A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, 4, 287.
- Chenu, K., Porter, J. R., Martre, P., Basso, B., Chapman, S. C., Ewert, F., Bindi, M. and Asseng, S. 2017. Contribution of Crop Models to Adaptation in Wheat. *Trends in Plant Science*, 22, 472-490.
- Cheuk, A. and Houde, M. 2018. A New Barley Stripe Mosaic Virus Allows Large Protein Overexpression for Rapid Function Analysis. *Plant Physiology*, 176, 1919.
- Christopher, J. T., Christopher, M. J., Borrell, A. K., Fletcher, S. and Chenu, K. 2016. Stay-green traits to improve wheat adaptation in well-watered and water-limited environments. *Journal of Experimental Botany*, 67, 5159-5172.
- Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., Borrill, P., Kettleborough, G., Heavens, D., Chapman, H., Lipscombe, J., Barker, T., Lu, F.-H., Mckenzie, N., Raats, D., Ramirez-Gonzalez, R. H., Coince, A., Peel, N., Percival-Alwyn, L., Duncan, O., Trösch, J., Yu, G., Bolser, D. M., Namaati, G., Kerhornou, A., Spannagl, M., Gundlach, H., Haberer, G., Davey, R. P., Fosker, C., Palma, F. D., Phillips, A., Millar, A. H., Kersey, P. J., Uauy, C., Krasileva, K. V., Swarbreck, D., Bevan, M. W. and Clark, M. D. 2017. An improved

- assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research*, 27, 885-896.
- Cossani, C. M. and Reynolds, M. P. 2015. Heat Stress Adaptation in Elite Lines Derived from Synthetic Hexaploid Wheat. *Crop Science*, 55, 2719-2735.
- Crain, J., Mondal, S., Rutkoski, J., Singh, R. P. and Poland, J. 2018. Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *The Plant Genome*, 11.
- Demirer, G. S., Zhang, H., Matos, J., Goh, N., Cunningham, F. J., Sung, Y., Chang, R., Aditham, A. J., Chio, L., Cho, M.-J., Staskawicz, B. and Landry, M. P. 2018. High Aspect Ratio Nanomaterials Enable Delivery of Functional Genetic Material Without DNA Integration in Mature Plants. *bioRxiv*, 10.1101/179549.
- Dixon, L. E., Greenwood, J. R., Bencivenga, S., Zhang, P., Cockram, J., Mellers, G., Ramm, K., Cavanagh, C., Swain, S. M. and Boden, S. A. 2018. *TEOSINTE BRANCHED 1* Regulates Inflorescence Architecture and Development in Bread Wheat (*Triticum aestivum* L.). *The Plant Cell*, 30, 563-581.
- Dreisigacker, S., Kishii, M., Lage, J. and Warburton, M. 2008. Use of synthetic hexaploid wheat to increase diversity for CIMMYT bread wheat improvement. *Australian Journal of Agricultural Research*, 59, 413-420.
- Dubcovsky, J. and Dvorak, J. 2007. Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication. *Science*, 316, 1862.
- Dvorak, J., Akhunov, E. D., Akhunov, A. R., Deal, K. R. and Luo, M.-C. 2006. Molecular Characterization of a Diagnostic DNA Marker for Domesticated Tetraploid Wheat Provides Evidence for Gene Flow from Wild Tetraploid Wheat to Hexaploid Wheat. *Molecular Biology and Evolution*, 23, 1386-1396.
- Fao, I. a. W. 2015. The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome: FAO.
- Furbank, R. T. and Tester, M. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, 16, 635-644.
- Gallusci, P., Dai, Z., Génard, M., Gauffretau, A., Leblanc-Fournier, N., Richard-Molard, C., Vile, D. and Brunel-Muguet, S. 2017. Epigenetics for Plant Improvement: Current Knowledge and Modeling Avenues. *Trends in Plant Science*, 22, 610-623.
- Gardiner, L.-J., Bansept-Basler, P., Olohan, L., Joynson, R., Brenchley, R., Hall, N., O'sullivan, D. M. and Hall, A. 2016. Mapping-by-sequencing in complex polyploid genomes using genic sequence capture: a case study to map yellow rust resistance in hexaploid wheat. *The Plant Journal*, 87, 403-419.
- Gardiner, L.-J., Brabbs, T., Akhunova, A., Jordan, K., Budak, H., Richmond, T., Singh, S., Catchpole, L., Akhunov, E. and Hall, A. 2018a. Integrating genomic resources to present full gene and promoter capture probe sets for bread wheat. *bioRxiv*, 10.1101/363663.
- Gardiner, L.-J., Joynson, R., Omony, J., Rusholme-Pilcher, R., Olohan, L., Lang, D., Bai, C., Hawkesford, M., Salt, D., Spannagl, M., Mayer, K. F. X., Kenny, J., Bevan, M., Hall, N. and Hall, A. 2018b. Hidden variation in polyploid wheat drives local adaptation. *Genome Research*, 28, 1319-1332.
- Gaut, B. S., Seymour, D. K., Liu, Q. and Zhou, Y. 2018. Demography and its effects on genomic variation in crop domestication. *Nature Plants*, 4, 512-520.
- Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C. and Lucretti, S. 2013. FISHIS: Fluorescence In Situ Hybridization in Suspension and Chromosome Flow Sorting Made Easy. *PLOS ONE*, 8, e57994.
- Golan, G., Oksenberg, A. and Peleg, Z. 2015. Genetic evidence for differential selection of grain and embryo weight during wheat evolution under domestication. *Journal of Experimental Botany*, 66, 5703-5711.

- Golicz, A. A., Bayer, P. E. and Edwards, D. 2015. Skim-Based Genotyping by Sequencing. In: BATLEY, J. (ed.) *Plant Genotyping: Methods and Protocols*. New York, NY: Springer New York.
- Grinberg, N. F., Orhobor, O. I. and King, R. D. 2018. An Evaluation of Machine-learning for Predicting Phenotype: Studies in Yeast, Rice and Wheat. *bioRxiv*, 10.1101/105528.
- Halloran, G. M., Ogonnaya, F. C. and Lagudah, E. S. 2008. *Triticum (Aegilops) tauschii* in the natural and artificial synthesis of hexaploid wheat. *Australian Journal of Agricultural Research*, 59, 475-490.
- Harwood, W. A. 2012. Advances and remaining challenges in the transformation of barley and wheat. *Journal of Experimental Botany*, 63, 1791-1798.
- Hassani-Pak, K., Castellote, M., Esch, M., Hindle, M., Lysenko, A., Taubert, J. and Rawlings, C. 2016. Developing integrated crop knowledge networks to advance candidate gene discovery. *Applied & Translational Genomics*, 11, 18-26.
- Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S. and David, J. 2007. Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication. *Molecular Biology and Evolution*, 24, 1506-1517.
- Hawkesford, M. J., Araus, J.-L., Park, R., Calderini, D., Miralles, D., Shen, T., Zhang, J. and Parry, M. a. J. 2013. Prospects of doubling global wheat yields. *Food and Energy Security*, 2, 34-48.
- Hen-Avivi, S., Savin, O., Racovita, R., Lee, W., Adamski, N., Malitsky, S., Almekias-Siegl, E., Levy, M., Vautrin, S., Bergès, H., Friedlander, G., Kartvelishvili, E., Ben-Zvi, G., Alkan, N., Uauy, C., Kanyuka, K., Jetter, R., Distelfeld, A. and Aharoni, A. 2016. A Metabolic Gene Cluster in the Wheat *W1* and the Barley *Cer-cqu* Loci Determines β -Diketone Biosynthesis and Glauconess. *The Plant Cell*, 28, 1440-1460.
- Higgins, J. A., Bailey, P. C. and Laurie, D. A. 2010. Comparative Genomics of Flowering Time Pathways Using *Brachypodium distachyon* as a Model for the Temperate Grasses. *PLOS ONE*, 5, e10065.
- Himi, E., Maekawa, M., Miura, H. and Noda, K. 2011. Development of PCR markers for *Tamyb10* related to *R-1*, red grain color gene in wheat. *Theoretical and Applied Genetics*, 122, 1561-1576.
- Huang, B. E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K. and Cavanagh, C. R. 2012. A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnology Journal*, 10, 826-839.
- Huang, X. and Han, B. 2014. Natural Variations and Genome-Wide Association Studies in Crop Plants. *Annual Review of Plant Biology*, 65, 531-551.
- Ishida, Y., Tsunashima, M., Hiei, Y. and Komari, T. 2015. Wheat (*Triticum aestivum* L.) Transformation Using Immature Embryos. In: WANG, K. (ed.) *Agrobacterium Protocols: Volume 1*. New York, NY: Springer New York.
- Iwgc 2014. A chromosome-based draft sequence of the hexaploid bread wheat *Triticum aestivum* genome. *Science*, 345, 1251788.
- Iwgc, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Pozniak, C. J., Stein, N., Choulet, F., Distelfeld, A., Eversole, K., Poland, J., Rogers, J., Ronen, G., Sharpe, A. G., Pozniak, C., Ronen, G., Stein, N., Barad, O., Baruch, K., Choulet, F., Keeble-Gagnère, G., Mascher, M., Sharpe, A. G., Ben-Zvi, G., Josselin, A.-A., Stein, N., Mascher, M., Himmelbach, A., Choulet, F., Keeble-Gagnère, G., Mascher, M., Rogers, J., Balfourier, F., Gutierrez-Gonzalez, J., Hayden, M., Josselin, A.-A., Koh, C., Muehlbauer, G., Pasam, R. K., Paux, E., Pozniak, C. J., Rigault, P., Sharpe, A. G., Tibbits, J., Tiwari, V., Choulet, F., Keeble-Gagnère, G., Mascher, M., Josselin, A.-A., Rogers, J., Spannagl, M., Choulet, F., Lang, D., Gundlach, H., Haberer, G., Keeble-Gagnère, G., Mayer, K. F. X., Ormanbekova, D., Paux, E., Prade, V., Šimková, H., Wicker, T., Choulet, F., Spannagl, M., Swarbrick, D., Rimbart, H., Felder, M., Guilhot, N., Gundlach, H., Haberer, G., Kaithakottil, G., Keilwagen, J., Lang, D., Leroy, P., Lux, T., Mayer, K. F. X., Twardziok, S., Venturini, L., Appels, R., Rimbart, H., Choulet, F., Juhász, A., Keeble-Gagnère, G., Choulet, F., Spannagl, M., Lang, D., Abrouk, M., Haberer, G., Keeble-

- Gagnère, G., Mayer, K. F. X., Wicker, T., Choulet, F., Wicker, T., Gundlach, H., Lang, D., Spannagl, M., Lang, D., Spannagl, M., et al. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361, eaar7191.
- Jordan, K. W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P., Sherman, J., Akhunova, A., Blake, N. K., Pumphrey, M. O., Glover, K., Dubcovsky, J., Talbert, L. and Akhunov, E. D. 2018. The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *The Plant Journal*, 95, 1039-1054.
- Jordan, K. W., Wang, S., Lun, Y., Gardiner, L.-J., Maclachlan, R., Hucl, P., Wiebe, K., Wong, D., Forrest, K. L., Sharpe, A. G., Sidebottom, C. H. D., Hall, N., Toomajian, C., Close, T., Dubcovsky, J., Akhunova, A., Talbert, L., Bansal, U. K., Bariana, H. S., Hayden, M. J., Pozniak, C., Jeddelloh, J. A., Hall, A., Akhunov, E. and Consortium, I. 2015. A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biology*, 16, 48.
- Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I. and Jones, J. D. G. 2013. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal*, 76, 530-544.
- Keller, B., Wicker, T. and Krattinger, S. G. 2016. Advances in Wheat and Pathogen Genomics: Implications for Disease Control. *Annual Review of Phytopathology*, 56, 67-87.
- Kim, W., Johnson, J. W., Graybosch, R. A. and Gaines, C. S. 2003. Physicochemical Properties and End-use Quality of Wheat Starch as a Function of Waxy Protein Alleles. *Journal of Cereal Science*, 37, 195-204.
- King, J., Grewal, S., Yang, C.-Y., Hubbart, S., Scholefield, D., Ashling, S., Edwards, K. J., Allen, A. M., Burridge, A., Bloor, C., Davassi, A., Silva, G. J., Chalmers, K. and King, I. P. 2016. A step change in the transfer of interspecific variation into wheat from *Amblyopyrum muticum*. *Plant Biotechnology Journal*, 15, 217-226.
- Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., Simmonds, J., Ramirez-Gonzalez, R. H., Wang, X., Borrill, P., Fosker, C., Ayling, S., Phillips, A. L., Uauy, C. and Dubcovsky, J. 2017. Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences*, 114, E913.
- Kugler, K. G., Siegwart, G., Nussbaumer, T., Ametz, C., Spannagl, M., Steiner, B., Lemmens, M., Mayer, K. F. X., Buerstmayr, H. and Schweiger, W. 2013. Quantitative trait loci-dependent analysis of a gene co-expression network associated with Fusarium head blight resistance in bread wheat (*Triticum aestivum* L.). *BMC Genomics*, 14, 728.
- Lee, T., Hwang, S., Kim, C. Y., Shim, H., Kim, H., Ronald, P. C., Marcotte, E. M. and Lee, I. 2017. WheatNet: a Genome-Scale Functional Network for Hexaploid Bread Wheat, *Triticum aestivum*. *Molecular Plant*, 10, 1133-1136.
- Lee, W.-S., Rudd, J. J. and Kanyuka, K. 2015. Virus induced gene silencing (VIGS) for functional analysis of wheat genes involved in *Zymoseptoria tritici* susceptibility and resistance. *Fungal Genetics and Biology*, 79, 84-88.
- Li, W. and Yang, B. 2017. Translational genomics of grain size regulation in wheat. *Theoretical and Applied Genetics*, 130, 1765-1771.
- Ling, H.-Q., Ma, B., Shi, X., Liu, H., Dong, L., Sun, H., Cao, Y., Gao, Q., Zheng, S., Li, Y., Yu, Y., Du, H., Qi, M., Li, Y., Lu, H., Yu, H., Cui, Y., Wang, N., Chen, C., Wu, H., Zhao, Y., Zhang, J., Li, Y., Zhou, W., Zhang, B., Hu, W., Van Eijk, M. J. T., Tang, J., Witsenboer, H. M. A., Zhao, S., Li, Z., Zhang, A., Wang, D. and Liang, C. 2018. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. *Nature*, 557, 424-428.
- Liu, W., Maccaferri, M., Bulli, P., Rynearson, S., Tuberosa, R., Chen, X. and Pumphrey, M. 2017. Genome-wide association mapping for seedling and field resistance to *Puccinia striiformis* f. sp. *tritici* in elite durum wheat. *Theoretical and Applied Genetics*, 130, 649-667.

- Liu, Z., Xin, M., Qin, J., Peng, H., Ni, Z., Yao, Y. and Sun, Q. 2015. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biology*, 15, 152.
- Lowe, A., Harrison, N. and French, A. P. 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, 13, 80.
- Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.-J., Scelonge, C., Lenderts, B., Chamberlin, M., Cushatt, J., Wang, L., Ryan, L., Khan, T., Chow-Yiu, J., Hua, W., Yu, M., Banh, J., Bao, Z., Brink, K., Igo, E., Rudrappa, B., Shamseer, P. M., Bruce, W., Newman, L., Shen, B., Zheng, P., Bidney, D., Falco, C., Register, J., Zhao, Z.-Y., Xu, D., Jones, T. and Gordon-Kamm, W. 2016. Morphogenic Regulators *Baby boom* and *Wuschel* Improve Monocot Transformation. *The Plant Cell*, 28, 1998.
- Mackay, I. J., Bansept-Basler, P., Barber, T., Bentley, A. R., Cockram, J., Gosman, N., Greenland, A. J., Horsnell, R., Howells, R., O'sullivan, D. M., Rose, G. A. and Howell, P. J. 2014. An Eight-Parent Multiparent Advanced Generation Inter-Cross Population for Winter-Sown Wheat: Creation, Properties, and Validation. *G3: Genes/Genomes/Genetics*, 4, 1603.
- Maher, K. A., Bajic, M., Kajala, K., Reynoso, M., Pauluzzi, G., West, D. A., Zumstein, K., Woodhouse, M., Bubb, K., Dorrity, M. W., Queitsch, C., Bailey-Serres, J., Sinha, N., Brady, S. M. and Deal, R. B. 2018. Profiling of Accessible Chromatin Regions across Multiple Plant Species and Cell Types Reveals Common Gene Regulatory Principles and New Control Modules. *The Plant Cell*, 30, 15.
- Mairhofer, S., Sturrock, C. J., Bennett, M. J., Mooney, S. J. and Pridmore, T. P. 2015. Extracting multiple interacting root systems using X-ray microcomputed tomography. *The Plant Journal*, 84, 1034-1043.
- Marchal, C., Zhang, J., Zhang, P., Fenwick, P., Steuernagel, B., Adamski, N. M., Boyd, L., McIntosh, R., Wulff, B. B. H., Berry, S., Lagudah, E. and Uauy, C. 2018. BED-domain-containing immune receptors confer diverse resistance spectra to yellow rust. *Nature Plants*, 4, 662-668.
- Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutayavin, T., Stehling-Sun, S., Johnson, A. K., Canfield, T. K., Giste, E., Diegel, M., Bates, D., Hansen, R. S., Neph, S., Sabo, P. J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S. R., Kaul, R. and Stamatoyannopoulos, J. A. 2012. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science (New York, N.Y.)*, 337, 1190-1195.
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and Depristo, M. A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297-1303.
- Milner, S. G., Maccaferri, M., Huang, B. E., Mantovani, P., Massi, A., Frascaroli, E., Tuberosa, R. and Salvi, S. 2015. A multiparental cross population for mapping QTL for agronomic traits in durum wheat (*Triticum turgidum* ssp. *durum*). *Plant Biotechnology Journal*, 14, 735-748.
- Mo, Y., Howell, T., Vasquez-Gross, H., De Haro, L. A., Dubcovsky, J. and Pearce, S. 2018. Mapping causal mutations by exome sequencing in a wheat TILLING population: a tall mutant case study. *Molecular Genetics and Genomics*, 293, 463-477.
- Muterko, A., Kalendar, R., Cockram, J. and Balashova, I. 2015. Discovery, evaluation and distribution of haplotypes and new alleles of the *Photoperiod-A1* gene in wheat. *Plant Molecular Biology*, 88, 149-164.
- Nachmanson, D., Lian, S., Schmidt, E. K., Hipp, M. J., Baker, K. T., Zhang, Y., Tretiakova, M., Loubet-Senear, K., Kohn, B. F., Salk, J. J., Kennedy, S. R. and Risques, R. A. 2018. Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). *Genome Research*, 28, 1589-1599.

- Nakamura, S., Pourkheirandish, M., Morishige, H., Kubo, Y., Nakamura, M., Ichimura, K., Seo, S., Kanamori, H., Wu, J., Ando, T., Hensel, G., Sameri, M., Stein, N., Sato, K., Matsumoto, T., Yano, M. and Komatsuda, T. 2016. Mitogen-Activated Protein Kinase Kinase 3 Regulates Seed Dormancy in Barley. *Current Biology*, 26, 775-781.
- Parent, B., Shahinnia, F., Maphosa, L., Berger, B., Rabie, H., Chalmers, K., Kovalchuk, A., Langridge, P. and Fleury, D. 2015. Combining field performance with controlled environment plant imaging to identify the genetic control of growth and transpiration underlying yield response to water-deficit stress in wheat. *Journal of Experimental Botany*, 66, 5481-5492.
- Pearce, S., Kippes, N., Chen, A., Debernardi, J. M. and Dubcovsky, J. 2016. RNA-seq studies using wheat *PHYTOCHROME B* and *PHYTOCHROME C* mutants reveal shared and specific functions in the regulation of flowering and shade-avoidance pathways. *BMC Plant Biology*, 16, 141.
- Pearce, S., Tabbita, F., Cantu, D., Buffalo, V., Avni, R., Vazquez-Gross, H., Zhao, R., Conley, C. J., Distelfeld, A. and Dubcovsky, J. 2014. Regulation of Zn and Fe transporters by the *GPC1* gene during early wheat monocarpic senescence. *BMC Plant Biology*, 14, 368.
- Perez-Sanz, F., Navarro, P. J. and Egea-Cortines, M. 2017. Plant phenomics: an overview of image acquisition technologies and image data analysis algorithms. *GigaScience*, 6, gix092-gix092.
- Periyannan, S. 2017. Sustaining global agriculture through rapid detection and deployment of genetic resistance to deadly crop diseases. *New Phytologist*, 219, 45-51.
- Pfeifer, J., Kirchgessner, N., Colombi, T. and Walter, A. 2015. Rapid phenotyping of crop root systems in undisturbed field soils using X-ray computed tomography. *Plant Methods*, 11, 41.
- Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., Putten, W. H., Kleyer, M., Schurr, U. and Postma, J. 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist*, 212, 838-855.
- Ramegowda, V., Mysore, K. S. and Senthil-Kumar, M. 2014. Virus-induced gene silencing is a versatile tool for unraveling the functional relevance of multiple abiotic-stress-responsive genes in crop plants. *Frontiers in Plant Science*, 5, 323.
- Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., Van Ex, F., Pasha, A., Khedikar, Y., Robinson, S. J., Cory, A. T., Florio, T., Concia, L., Juery, C., Schoonbeek, H., Steuernagel, B., Xiang, D., Ridout, C. J., Chalhou, B., Mayer, K. F. X., Benhamed, M., Latrasse, D., Bendahmane, A., Wulff, B. B. H., Appels, R., Tiwari, V., Datla, R., Choulet, F., Pozniak, C. J., Provart, N. J., Sharpe, A. G., Paux, E., Spannagl, M., Bräutigam, A. and Uauy, C. 2018. The transcriptional landscape of polyploid wheat. *Science*, 361, eaar6089.
- Ramirez-Gonzalez, R. H., Uauy, C. and Caccamo, M. 2015. PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics*, 31, 2038-2039.
- Ray, D. K., Mueller, N. D., West, P. C. and Foley, J. A. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLOS ONE*, 8, e66428.
- Rey, M.-D., Martín, A. C., Higgins, J., Swarbreck, D., Uauy, C., Shaw, P. and Moore, G. 2017. Exploiting the *ZIP4* homologue within the wheat *Ph1* locus has identified two lines exhibiting homoeologous crossover in wheat-wild relative hybrids. *Molecular Breeding*, 37, 95.
- Rey, M.-D., Martín, A. C., Smedley, M., Hayta, S., Harwood, W., Shaw, P. and Moore, G. 2018. Magnesium Increases Homoeologous Crossover Frequency During Meiosis in *ZIP4* (*Ph1* Gene) Mutant Wheat-Wild Relative Hybrids. *Frontiers in Plant Science*, 9, 509.
- Reynolds, M. and Langridge, P. 2016. Physiological breeding. *Current Opinion in Plant Biology*, 31, 162-171.
- Richard, C., Hickey, L., Fletcher, S., Chenu, K., Borrell, A. and Christopher, J. 2015. High-throughput Phenotyping of Wheat Seminal Root Traits in a Breeding Context. *Procedia Environmental Sciences*, 29, 102-103.
- Richardson, T., Thistleton, J., Higgins, T. J., Howitt, C. and Ayliffe, M. 2014. Efficient *Agrobacterium* transformation of elite wheat germplasm without selection. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 119, 647-659.

- Rodgers-Melnick, E., Vera, D. L., Bass, H. W. and Buckler, E. S. 2016. Open chromatin reveals the functional maize genome. *Proceedings of the National Academy of Sciences*, 113, E3177.
- Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. and Lippman, Z. B. 2017. Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell*, 171, 470-480.e8.
- Rosegrant, M. W., Tokgoz, S., Bhandary, P. and Msangi, S. 2013. Looking ahead: Scenarios for the future of food. *2012 Global Food Policy Report*. Washington, D.C.: International Food Policy Research Institute.
- Saintenac, C., Jiang, D. and Akhunov, E. D. 2011. Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome. *Genome Biology*, 12, R88.
- Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J. J., King, R. C., Marande, W., Powers, S. J., Bergès, H., Phillips, A. L., Uauy, C., Hammond-Kosack, K. E., Langin, T. and Kanyuka, K. 2018. Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. *Nature Genetics*, 50, 368-374.
- Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., Vrána, J., Kubaláková, M., Krattinger, S. G., Wicker, T., Doležel, J., Keller, B. and Wulff, B. B. H. 2016. Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biology*, 17, 221.
- Schneider, L. M., Adamski, N. M., Christensen, C. E., Stuart, D. B., Vautrin, S., Hansson, M., Uauy, C. and Von Wettstein-Knowles, P. 2016. The *Cer-cqu* gene cluster determines three key players in a β -diketone synthase polyketide pathway synthesizing aliphatics in epicuticular waxes. *Journal of Experimental Botany*, 67, 2715-2730.
- Shorinola, O., Balcáková, B., Hyles, J., Tibbits, J. F. G., Hayden, M. J., Holuřova, K., Valárik, M., Distelfeld, A., Torada, A., Barrero, J. M. and Uauy, C. 2017. Haplotype Analysis of the Pre-harvest Sprouting Resistance Locus *Phs-A1* Reveals a Causal Role of *TaMKK3-A* in Global Germplasm. *Frontiers in Plant Science*, 8, 1555.
- Shorinola, O., Kaye, R., Kepinski, S. and Uauy, C. 2018. Isolation and characterization of mutants with altered seminal root numbers in hexaploid wheat. *bioRxiv*, 10.1101/364018.
- Silva-Perez, V., Molero, G., Serbin, S. P., Condon, A. G., Reynolds, M. P., Furbank, R. T. and Evans, J. R. 2018. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *Journal of Experimental Botany*, 69, 483-496.
- Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., Del Blanco, A., Dubcovsky, J. and Uauy, C. 2016. A splice acceptor site mutation in *TaGW2-A1* increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theoretical and Applied Genetics*, 129, 1099-1112.
- Singh, A., Ganapathysubramanian, B., Singh, A. K. and Sarkar, S. 2016. Machine Learning for High-Throughput Stress Phenotyping in Plants. *Trends in Plant Science*, 21, 110-124.
- Singh, N., Wu, S., Raupp, J., Sehgal, S., Arora, S., Tiwari, V., Vikram, P., Singh, S., Chhuneja, P., Gill, B. and Poland, J. 2018. Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. *bioRxiv*, 10.1101/410779.
- Sirault, X. R. R., James, R. A. and Furbank, R. T. 2009. A new screening method for osmotic component of salinity tolerance in cereals using infrared thermography. *Functional Plant Biology*, 36, 970-977.
- Slade, A. J., Mcguire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K. M., Steine, M. N., Goodstal, J. F. and Knauf, V. C. 2012. Development of high amylose wheat through TILLING. *BMC Plant Biology*, 12, 69.
- Spieritz, H. 2012. Avenues to meet food security. The role of agronomy on solving complexity in food production and resource use. *European Journal of Agronomy*, 43, 1-8.
- Springer, N. M. 2013. Epigenetics and crop improvement. *Trends in Genetics*, 29, 241-247.
- Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., Hatta, A., Ayliffe, M., Bariana, H., Jones, J. D. G., Lagudah, E. S. and Wulff, B. B. H. 2016. Rapid cloning

- of disease-resistance genes in plants using mutagenesis and sequence capture. *Nature Biotechnology*, 34, 652.
- Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P. and Reynolds, M. P. 2015. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theoretical and Applied Genetics*, 128, 353-363.
- Sullivan, Alessandra m., Arsovski, Andrej a., Lempe, J., Bubb, Kerry l., Weirauch, Matthew t., Sabo, Peter j., Sandstrom, R., Thurman, Robert e., Neph, S., Reynolds, Alex p., Stergachis, Andrew b., Vernet, B., Johnson, Audra k., Haugen, E., Sullivan, Shawn t., Thompson, A., Neri, Fidencio v., Weaver, M., Diegel, M., Mnaimneh, S., Yang, A., Hughes, Timothy r., Nemhauser, Jennifer l., Queitsch, C. and Stamatoyannopoulos, John a. 2014. Mapping and Dynamics of Regulatory DNA and Transcription Factor Networks in *A. thaliana*. *Cell Reports*, 8, 2015-2030.
- Sutherland, B. L. and Galloway, L. F. 2016. Postzygotic isolation varies by ploidy level within a polyploid complex. *New Phytologist*, 213, 404-412.
- Thind, A. K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J. and Krattinger, S. G. 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nature Biotechnology*, 35, 793.
- Tilman, D., Balzer, C., Hill, J. and Befort, B. L. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108, 20260.
- Torada, A., Koike, M., Ogawa, T., Takenouchi, Y., Tadamura, K., Wu, J., Matsumoto, T., Kawaura, K. and Ogihara, Y. 2016. A Causal Gene for Seed Dormancy on Wheat Chromosome 4A Encodes a MAP Kinase Kinase. *Current Biology*, 26, 782-787.
- Trachsel, S., Kaeppler, S. M., Brown, K. M. and Lynch, J. P. 2011. Shovelomics: high throughput phenotyping of maize (*Zea mays* L.) root architecture in the field. *Plant and Soil*, 341, 75-87.
- Tsaftaris, S. A., Minervini, M. and Scharr, H. 2016. Machine Learning for Plant Phenotyping Needs Image Processing. *Trends in Plant Science*, 21, 989-991.
- Turner, A., Beales, J., Faure, S., Dunford, R. P. and Laurie, D. A. 2005. The Pseudo-Response Regulator *Ppd-H1* Provides Adaptation to Photoperiod in Barley. *Science*, 310, 1031.
- Uauy, C. 2017. Wheat genomics comes of age. *Current Opinion in Plant Biology*, 36, 142-148.
- Uauy, C., Brevis, J. C., Chen, X., Khan, I., Jackson, L., Chicaiza, O., Distelfeld, A., Fahima, T. and Dubcovsky, J. 2005. High-temperature adult-plant (HTAP) stripe rust resistance gene *Yr36* from *Triticum turgidum* ssp. *dicoccoides* is closely linked to the grain protein content locus *Gpc-B1*. *Theoretical and Applied Genetics*, 112, 97.
- Uauy, C., Brevis, J. C. and Dubcovsky, J. 2006. The high grain protein content gene *Gpc-B1* accelerates senescence and has pleiotropic effects on protein content in wheat. *Journal of Experimental Botany*, 57, 2785-2794.
- Uauy, C., Wulff, B. B. H. and Dubcovsky, J. 2017. Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat. *Annual Review of Genetics*, 51, 435-454.
- Un. 2015. *The World Population Prospects: 2015 Revision* [Online]. Available: <http://www.un.org/en/development/desa/publications/world-population-prospects-2015-revision.html> [Accessed].
- Voss-Fels, K. P., Qian, L., Parra-Londono, S., Uptmoor, R., Frisch, M., Keeble-Gagnère, G., Appels, R. and Snowdon, R. J. 2016. Linkage drag constrains the roots of modern wheat. *Plant, Cell & Environment*, 40, 717-725.
- Wang, D. R., Agosto-Pérez, F. J., Chebotarov, D., Shi, Y., Marchini, J., Fitzgerald, M., McNally, K. L., Alexandrov, N. and Mccouch, S. R. 2018a. An imputation platform to enhance integration of rice genetic resources. *Nature Communications*, 9, 3519.
- Wang, K., Liu, H., Du, L. and Ye, X. 2016. Generation of marker-free transgenic hexaploid wheat via an *Agrobacterium*-mediated co-transformation strategy in commercial Chinese wheat varieties. *Plant Biotechnology Journal*, 15, 614-623.

- Wang, W., Simmonds, J., Pan, Q., Davidson, D., He, F., Battal, A., Akhunova, A., Trick, H. N., Uauy, C. and Akhunov, E. 2018b. Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of *TaGW2* homoeologues to grain size and weight in wheat. *Theoretical and Applied Genetics*, 131, 2463-2475.
- Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.-L. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, 32, 947.
- Wang, Y., Yu, H., Tian, C., Sajjad, M., Gao, C., Tong, Y., Wang, X. and Jiao, Y. 2017. Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiology*, 175, 746-757.
- Wilhelm, E. P., Turner, A. S. and Laurie, D. A. 2009. Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.). *Theoretical and Applied Genetics*, 118, 285-294.
- Wilkinson, P. A., Winfield, M. O., Barker, G. L. A., Tyrrell, S., Bian, X., Allen, A. M., BurrIDGE, A., Coghill, J. A., Waterfall, C., Caccamo, M., Davey, R. P. and Edwards, K. J. 2016. CerealsDB 3.0: expansion of resources and data integration. *BMC Bioinformatics*, 17, 256.
- Winfield, M. O., Allen, A. M., BurrIDGE, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., King, J., West, C., Griffiths, S., King, I., Bentley, A. R. and Edwards, K. J. 2015. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnology Journal*, 14, 1195-1206.
- Winfield, M. O., Allen, A. M., Wilkinson, P. A., BurrIDGE, A. J., Barker, G. L. A., Coghill, J., Waterfall, C., Wingen, L. U., Griffiths, S. and Edwards, K. J. 2017. High-density genotyping of the A.E. Watkins Collection of hexaploid landraces identifies a large molecular diversity compared to elite bread wheat. *Plant Biotechnology Journal*, 16, 165-175.
- Wingen, L. U., Orford, S., Goram, R., Leverington-Waite, M., Bilham, L., Patsiou, T. S., Ambrose, M., Dicks, J. and Griffiths, S. 2014. Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat. *Theoretical and Applied Genetics*, 127, 1831-1842.
- Wingen, L. U., West, C., Leverington-Waite, M., Collier, S., Orford, S., Goram, R., Yang, C.-Y., King, J., Allen, A. M., BurrIDGE, A., Edwards, K. J. and Griffiths, S. 2017. Wheat Landrace Genome Diversity. *Genetics*, 205, 1657-1676.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V. and Provart, N. J. 2007. An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLOS ONE*, 2, e718.
- Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., McMullen, M. D. and Gaut, B. S. 2005. The Effects of Artificial Selection on the Maize Genome. *Science*, 308, 1310.
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T. and Dubcovsky, J. 2003. Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences*, 100, 6263.
- York, L. M., Slack, S., Bennett, M. J. and Foulkes, M. J. 2018. Wheat shovelomics I: A field phenotyping approach for characterising the structure and function of root systems in tillering species. *bioRxiv*, 10.1101/280875.
- Zhang, H., Yang, Y., Wang, C., Liu, M., Li, H., Fu, Y., Wang, Y., Nie, Y., Liu, X. and Ji, W. 2014. Large-scale transcriptome comparison reveals distinct gene activations in wheat responding to stripe rust and powdery mildew. *BMC Genomics*, 15, 898.
- Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L., Liu, J., Chen, K., Zhang, H., Gao, C. and Wang, D. 2018. Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits. *The Plant Journal*, 94, 857-866.
- Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., Qiu, J.-L. and Gao, C. 2016. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nature Communications*, 7, 12617.

- Accepted Article
- Zhao, G., Zou, C., Li, K., Wang, K., Li, T., Gao, L., Zhang, X., Wang, H., Yang, Z., Liu, X., Jiang, W., Mao, L., Kong, X., Jiao, Y. and Jia, J. 2017. The *Aegilops tauschii* genome reveals multiple impacts of transposons. *Nature Plants*, 3, 946-955.
- Zimin, A. V., Puiu, D., Hall, R., Kingan, S., Clavijo, B. J. and Salzberg, S. L. 2017. The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *GigaScience*, 6, gix097-gix097.

Box 1: Glossary of terminology

Term	Definition
Breeding programme	The development of novel plant cultivars by the deliberate crossing of plants, followed by the selection of the best resulting progeny over several generations.
Cultivar	A homozygous wheat line which has been selectively bred and is cultivated.
Genetic mapping	A method to delimit the position of a trait or gene within the genome. This is generally achieved using genetic or phenotypic markers in conjunction with mapping populations segregating for the trait or gene of interest.
Haplotype	A co-inherited block of DNA containing sequence polymorphisms which in wheat often span several genes.
Homoeolog	The chromosomes in polyploid species which are derived from different ancestral species (see Figure 1) and cannot pair and recombine during meiosis. The genes within these non-recombining chromosomes are referred to as homoeologs: in wheat there are three homoeologs (A, B, and D) for most genes.
Landrace	A domesticated and locally-adapted wheat line which is typically grown from farmer-saved seed and has not been modified through a breeding programme.
Ortholog	Genes in different species which evolved from a common ancestral gene.
Pangenome	The entire spectrum of genetic variation within a species, including genes and other variation found only in a subset of cultivars.
Phenomics	The high-throughput study of phenotypes.
Phenotype	The physical characteristics of an organism.
Polyploid	An organism with more than two sets of homologous (pairing) chromosomes.

Qualitative trait

A categorical characteristic, such as the number of lateral roots, frequently determined by a single genetic locus.

Quantitative trait

A continuous characteristic, such as yield or grain size, frequently determined by multiple small-effect genetic loci.

SNP

A single nucleotide polymorphism (SNP) is a variant in a single nucleotide base of DNA, such as a guanine to adenosine (G to A) change.

Wild relative

Plant species which are closely related to a crop species, but are not themselves domesticated nor cultivated.

Figure legends

Figure 1: Polyploid wheat genetic diversity. Genetic diversity in wild relatives was lost through the initial domestication of the tetraploid wild emmer and the subsequent natural hybridization which gave rise to the first hexaploid wheat (~10,000 years ago). Inter-breeding of hexaploid and emmer wheat reintroduced genetic diversity to the A and B genomes of hexaploid wheat (shown by shaded bars) which is present in landraces, whereas the wild progenitor of the D genome (goat grass) was reproductively isolated due to the difference in ploidy levels. Modern day cultivars have further reduced genetic diversity due to the bottlenecks imposed by artificial selection (breeding). Circles represent genetic diversity from the A (green), B (purple), and D (orange) genomes. Shaded bars show flow of genetic diversity. Time advances from top to bottom.

Figure 2: Effects of polyploidy on phenotypic variation. Mutations in a single homoeolog can lead to gain-of-function alleles which dominate over the other gene copies (dominance). This variation has been strongly selected upon and includes quantitative trait loci (QTL). Loci displaying dosage effects and functional redundancy are more difficult to select given that changes in one homoeolog lead to subtle or no phenotypic effects. However, combining mutations in multiple homoeologs can uncover an expanded phenotypic spectrum. This constitutes hidden variation of agronomic significance.

Figure 3: The future of wheat genomics. (A) A new frontier for wheat genomics research will focus on the effects of chromatin structural variation on gene expression regulation. Techniques such as Hi-C can be used to identify topologically-associating domains (TADs), chromatin regions which preferentially interact with each other. Such TADs can affect gene regulation by, for example, bringing distal regulatory elements such as enhancers into proximity with the regulated promoter. Regions of open chromatin can be identified by techniques such as ATAC-seq, the data from which can then be parsed to identify putative binding sites for regulatory proteins such as transcription factors (TF). Alongside investigation of chromatin conformation, studies into histone modifications are also necessary to further investigate the impact of epigenetic control on trait variation. (B) The new reference sequence has facilitated the development of many resources and techniques for studying gene function and regulation. Induced variation has been coupled with exome capture and targeted gene-family enrichment sequencing to study gene function. The *in silico* TILLING database identifies mutations in genes of interest, while techniques like MutRenSeq leverage targeted sequencing to rapidly clone *R*-genes. New advances in gene-editing technologies facilitate specific genetic changes, ranging from small indels and specific base-pair changes through to complete gene insertions. Expression data in wheat is

now easily accessible through sites such as www.wheat-expression.com and the eFP browser. Gene networks based on RNA-Seq data will facilitate our understanding of gene regulatory pathways in wheat. The improved genome sequence will also aid the study of epigenetic variation in wheat, particularly DNA methylation in coding and non-coding regions. (C) The sequencing of the wheat genome has started to provide insights into genetic variation between cultivars. Haplotype variation within genes is already available and will become increasingly informative as haplotypes extend across larger intervals and recombination blocks are defined. Gene annotations across wheat cultivars and wild relatives will provide insight into allelic variation across the wheat pangenome. Comparisons of transposon diversity, structural, and copy number variation between cultivars will also become increasingly possible with improved genome assemblies. These multiple layers of variation data will facilitate hypothesis generation and gene discovery in wheat. See text for references.

Figure 4: Integrated workflow for gene discovery in wheat. The development of genetic and genomic resources has accelerated candidate gene discovery in wheat. For example, multi-parent advanced generation inter-cross (MAGIC) populations were used to identify the wheat *Teosinte Branched 1* (*TB1*) homolog as a major regulator of inflorescence architecture and development (Dixon et al., 2018); the use of mapping populations and full genome sequence of the landrace reference and cultivars facilitated the identification of a candidate for the *Zymoseptoria tritici* resistance gene *STB6* (Saintenac et al., 2018); homology with rice enabled identification of *Grain Width 2* (*GW2*) (Wang et al., 2018b). Starting from a defined candidate gene (purple gene model in centre), there are multiple strategies for gene validation available in wheat. These include expression datasets, natural variation in cultivars and landraces, *in silico* EMS-mutants, multiple transgenic validations including gene editing, and transient transformation systems such as virus-induced gene silencing (VIGS). Studies have combined these approaches to confirm the gene underlying the trait of agronomic interest. Arrows denote the paths used to identify and validate *TB1* (purple), *STB6* (blue), and *GW2* (green).







